# Information Retrieval via Social Media

Farleen Azrina Zamberi, Nor Hafifah Hassan Adli, Norhayati Hussin, Masitah Ahmad

# Information Retrieval via Social Media

## Farleen Azrina Zamberi, Nor Hafifah Hassan Adli, Norhayati Hussin, Masitah Ahmad

Faculty of Information Management, Universiti Teknologi MARA, UiTM Selangor, Malaysia

**ABSTRACT**
People may search for information for various purposes such as interview, research or travel. However, in the digital era, people will easily find information via the internet, but it is hard to find scholarly information that suits their needs. The number of information retrieval tools had increase and people can use either one which relevant to their needs. The examples of information tools are OPAC, Internet search engine; subject directory, online database and digital library. By using social media, the hashtag is used in retrieval information and dissemination and mostly used in social media such as Instagram, Facebook and Twitter.
**Keywords:** Information Management, Information Retrieval, Social Media, Hashtag

## INTRODUCTION
The extensive and advancement of the internet made more accessible for people to access, retrieve and obtain information. The internet can be excellent as it stores a lot of information, but it is also very important for people to know the right method in retrieval the right information. The amount of information accessible via the Internet and the range of topics covered have grown to mammoth proportions. An increasing number of ever-more sophisticated tools have been developed to aid the retrieval of information from the Internet. Because of the enormous volume of information, there are some disadvantages need to be aware such as information overload and the problem in identifying the good quality of information. With the correct way to retrieving the information, people can minimize the percentage of information overload from occurring.

## INFORMATION RETRIEVAL
People face many options in accessing and retrieval information from the web. Some methods are successful, but the majority of people encounter a problem in retrieval the information they save and keep as they keep the various format of information. Information Retrieval (IR) is the process by which a collection of data is represented, stored, and searched for knowledge discovery as a response to a user request (Roshdi & Roohparvar, 2015). Various stages initiate with representing data and ending with returning relevant information to the user is the process involves in information retrieval. Filtering, searching, matching and ranking operations are including in the intermediate stage. The

main goal of an information retrieval system (IRS) is to find relevant information or a document that satisfies user information needs. There are some general applications for information retrieval systems which are a digital library, search engines and media search. As nowadays people become advanced equivalent with the enhancement of technologies, people started to retrieve information via the internet and social media. The internet contains various information that hardly to identify the validity. By social media, people also started to retrieve the information which are contains in the same group or interest by using hashtag.

## FINDING INFORMATION ON THE WEB
The Internet is known as the centre and source of information. Nowadays, people prefer used internet in searching for information rather than referring to books or journals. The major disadvantage of using Internet-based resources is the problem of locating good quality information from within the vast amount of information available (Bond, 2004).  People prefer to retrieve information from the internet because they can find the latest information from various sources and authors. But the disadvantages retrieving information from the internet is in identifying the quality and reliable information. The majority of Internet-based information used is found on the Worldwide Web (WWW). It is rapid growth, search tools change, new ones are developed, and the pool of information available is growing more rapidly than anyone would have imagined when it was first developed.

## SOCIAL MEDIA
Social media are famous as the medium to share ideas and communicate between researches, students and also the public. Because of the advancement in technology, the social media now already exist in applications format and this will make people easier in accessing the social media. Today, the most popular and used social media are Facebook, Twitter and Instagram. Through social media, not only people can share information, facts, and photos. Social media is freely available; people need devices and internet connection. Social media, published to reach people of similar interests, is often organized for later retrieval and dissemination by hashtags.

Social media employ mobile and web-based technologies to create highly interactive platforms via which individuals and communities share, co-create, discuss, and modify user-generated content (Kietzmann, Hermkens, McCarthy & Silvestre 2011). In the article title Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media, the author stated that there are seven functional blocks of social media which are identity, conversations, sharing, presence, relationships, reputation and groups.

Micro-blogging is the new trends of millions of population share their information, experience and even their activities in their daily such as a picnic, attend a seminar and jogging. Micro-blogging is listed as popular activities among teenager for them to spend their time to. Means, they can spend up to 4-8 hours only to micro-blogging since the characteristic of the micro-blog itself is easy platform for communication and light-weight that enables users to share everything in their daily lives. In Malaysia, the top micro-blog that recently and popularly used by Malaysian is Facebook, Twitter and Instagram. Facebook is the place where they can share information in the form of text and images or the combination of both while Twitter and Instagram is the place where they share the information

in the form of text and images respectively. Indeed, micro-blog contains a significant amount of data which can transform it into information and the example can be taken to analyze the data is from the profile user and their interest that contained or included in the profile. The benefit from it is the analyst and researcher can find the real-time data while analyzing the data through micro-blog.

In the simplest sentence, micro-blog is an action of posting any digital content such as image, voice, video and text and can be called as "online community". The introduction of micro-blog is started to popular and used by among of friends is start form era 2000's and about 9/10 people are using micro-blog in our daily life. The concept of micro-blog basically, we need to add the other user as friend or follower to know their activities. The activities of micro-blog are faster than a traditional blog that requires more time for a blogger to interact with the respondent. For example, the notification button eases the user to monitor any action made by their friends or followers.

The topic or content that is usually updating by the user is new stories, routine activities, current news and chosen interest that like by the user. The person who has financial background will post with financial information in his micro-blog while the person who has a medical background will post with medical information in his micro-blog.  Micro-blogging tools facilitate easily sharing status messages either publicly or within a social network and as well as a platform to them to mining the data regarding the user's interest.

**HASHTAGS**

Hashtagging in social networks essentially enables users to create subject headings for their posts at hand. Hashtagging is essentially Web 2.0 because it allows patrons to add and change not only content (data) but also content describing content (metadata). Hashtags assist in the ability to retrieve information and to follow trends (Buarki & Alkhateeb, 2017). Web 2.0 encourages Web users to share and publish user-generated data. There are five categories of tag taxonomy. Two categories of context-based tags with the first used to describe the content of an object and the second to provide the context of an object in which it was created; attribute tags which are inherent to the attributes of an object; subjective tags which express user opinion; and organizational tags which identify personal information. The hashtag is a simply a phrase or keywords that is preceded by a symbol (#) and used by the micro-blogging community to create a thread of conversation about certain topics or ideas. The usage of hashtag will allow the categorization of topics, sharing ideas and inspire others to exchange the ideas. The hashtag is very popular in social media such as Twitter, Instagram, Pinterest, and Facebook.

The hashtag is one of the ways to classify messages according to the topic. The user can retrieve the information when dealing with the labelled topic. As per Zappavigna said, the hashtag has developed into a "community building linguistic activity". The user can find information easily with the existence of hashtag.  A hashtag can be classified depending on the intention of the user to use the hashtag. The hashtag is used when the users want to express their comment, praise or criticize when to rise the event or people and to promote brands and also to provide updates on breaking news item and bring them to the content that easily findable and searchable across the social web.

It seems that displaying hashtags on the screen during a program increases the number of dedicated tweets, and therefore viewers' Twitter activity overall, as well as their engagement with the program during commercial breaks (Paola 2015), which means the hashtag bring to the other users to watch the TV program thus the viewer of the TV program is increasing and I called it as "effective online word-of-mouth"

A hashtag can be divided into two categories which are Topical hashtag and Non-topical hashtag (Brun and Moe, 2013). The topical hashtag is where the hashtag name is built for long-standing term and represent of the name of event or occasion. For example, the government launches a program or initiative to stop smoking, so they will build a hashtag like #StopSmoking or #SmokingNotGood and along the program, the government will use the hashtag for long period while the non-topical hashtag built for only for temporary period and nobody bother with the hashtag since the hashtag built for emotive marker. For example, he felt angry and started to update the status with hashtag #FeelingAngry; the hashtag will cease only at this stage. Different from the topical hashtag, the user tends to click the hashtag to know more about the hashtag built since the nature of hashtag itself is clickable and bring you to the numerous update of hashtag clicked.

According to Gregory and Jerome (2015), the hashtag is made by participants and willingness of the user. If the user hashtag something in his micro-blog means he is willing to hashtag his update. The hashtag classifies messages, improves the search process, and allows the organization as well as individual to link messages to existing knowledge and action communities. By right, the hashtag is one way to user to collect the followers.

**Table 1.** *Hashtag Codes with Examples and Frequencies*

| Hashtag category w/ definition | Examples | Freq | % |
|---|---|---|---|
| **1. KNOWLEDGE & PUBLIC EDUCATION** | | 413 | 50.5% |
| **Medical condition**–hashtag denoting disease or medical condition | #diabetes, #Hemophilia, #psoriasis | (153) | (18.7%) |
| **Knowledge base**–health-related research, knowledge, education | #hearthealthy, #AsthmaAwareness, #ALSresearch | (231) | (28.2%) |
| **Policy**–health-related public policy, public policy issues | #SunshineAct, #DeviceTax, #HCCosts | (29) | (3.5%) |
| **2. EVENTS**–health-related event, conference, holiday | #WorldCancerDay, #WalktoCureArthritis, #ALSIceBucketChallenge | 158 | 19.3% |
| **3. VALUES AND GOALS**– organizational values or goals. Useful for reinforcing the organization's core values and ultimate strategic goals. | #StopDiabetes, #PatientAccess, #coloncanceradvocate | 74 | 9.0% |
| **4. BRANDING**–organization-specific hashtags, unique organization identifiers, hashtags noting one of the organization's program names | #ShowUpDifferently, #MerckOncology, #UHFscholars | 59 | 7.2% |
| **5. DIALOGIG**–"chat" and dialogue hashtags | #HeartChat, #HemoChat, #DidYouKnow? | 41 | 5.0% |
| **6. TIME AND PLACE**–any time or location hashtag | #PuertoRico, #Summer, #Capitol | 27 | 3.3% |
| **7. CALL-TO-ACTION**–hashtags asking audience to do something | #ShareForAwareness, #HugDontJudge, #Raise100K | 26 | 3.2% |
| **8. BUSINESS**–related to business issues, stocks, companies, etc. Captures a wide range of non-health, non-advocacy-related hashtags, used in the sample | #Biotech, #GM, #Stocks | 18 | 2.2% |
| | **Total** | **818** | |

*Note:* Frequencies are the number of times each hashtag type was used in the 510 of 1,000 randomly selected tweets that were original (i.e., were not themselves retweets, n=774) and contained a hashtag (n=510). A total of 818 hashtags were included in these 510 tweets; percentages indicate proportions relative to these 818 hashtags. Not shown is a *miscellaneous* category in which 2 hashtags (#FREE, #adoption) were placed.

Table 1

Table 1 shows that the category hashtag that usually built by the user and its frequency. We can see that the hashtag category that updates most is knowledge and public education which contribute to 50.5% of the hashtag category and followed by an event which contributes to 19.3% of the hashtag category. Hashtag listed above is the example for a topical hashtag that explains earlier. The advantages from the hashtag built are user able to connect with others about the same topic. In this scenario, we can see that the information tend to travel faster with the existence of hashtag (Eduaro 2017) and we able to analyze the trend among the micro-blog users.

**CONCLUSIONS**

Social media has made easier by providing around-the-clock free feedback and exchange of ideas, as well as concerns of involvement in promoting and organizing an event. With the skills and experience, people may find a lot of information. Hashtag connects people by sharing their same interest and thoughts, and the use of a hashtag depends on its popularity through posting, following followers of the account and survival time. It is recommended that besides using hashtag only through online, it is used in offline media such as television and newspaper. It was difficult to qualify the search results of a hashtag because hashtags are not standardized and cannot be filtered by relevance or popularity.

According to Sindur, Nirav and Chandni, they believe that the problem from taking the information thru twitter (they use Twitter as an example and can be used for any micro-blog platform), there are no restriction, no custody and control by the related agency to monitor the information posted by the user. Everybody can post information even though the information is incorrect, wrong and counterfeit which probably can mislead the reader.

**Acknowledgement**

**REFERENCES**

Bond, C. S. (2004). Web users' information retrieval methods and skills. *Online Information Review, 28*(4). 254-259. doi: 10.1108/14684520410553741

Buarki, H. & Alkhateeb, B. (2017). Use of hashtags to retrieve information on the web. *The Electronic Library, 36*(2), 286-304. doi: 10.1108/EL-01-2017-0011

Caleffi,. P. (2015). The 'hashtag': A new word or a new rule?, *SKASE Journal of Theoretical Linguistics [online], 12*(2), 46-59. Retrieved from http://www.skase.sk/Volumes/JTL28/pdf_doc/05.pdf

Kietzmann, J. H., Hermkens, K., McCarthy, I. P. & Silvestre, B. S. (2011). Social Media? Get serious! Understanding the functional building blocks of social media. *Business Horizons, 54*, 241-251. doi:10.106/j.bushor.2011.01.005

Patel, N., Bhatt, N. & Shah, C. (2017). A survey of information retrieval on microblog. *International Journal of Computer Applications*, *161*(3), 35-38. doi: 10.5120/ijca2017913164

Roshdi, A. & Roohparvar, A. (2015). Review: Information Retrieval Techniques and Applications. *International Journal of Computer Networks and Communications Security, 3*(9), 373–377. Retrieved from http://www.ijcncs.org/published/volume3/issue9/p3_3-9.pdf

Salazar, E. (2017). *Hashtags 2.0* - An annotated history of the hashtag and a window to its future, *Icono 14, 15*(2), 16-54. doi: 10.7195/ri14. v15i2.1091

Saxton. G.D., Niyirora. J. N. & Waters. R.D. (2015). #AdvocatingForChange: The strategic use of hashtags in social media advocacy. *Advances in Social Work*, *16*(1). 154-169

Yazdanifard, R., Khalid, W., Wan Fadzilah Wan Yusoff & Babaei, H. R. (2011). Social Networks and Microblogging;The Emerging Marketing Trends&Tools of the Twenty-first Century. *International Conference on Computer Communication and Management*, *5*. 577-581. Retrieved from https://www.researchgate.net/publication/229053574_Social_Networks_and_Microblogging_The_Emerging_Marketing_TrendsTools_of_the_Twenty-first_Century